STATISTICAL ANALYSIS PLAN (SAP)

FOR NON-INTERVENTIONAL STUDIES

[redacted]

Non-Interventional Study Protocol
B1781044

Cohort Study of Venous Thromboembolism and Other Clinical Endpoints Among Osteoporotic Women Prescribed Bazedoxifene, Bisphosphonates or Raloxifene in Europe

Statistical Analysis Plan
(SAP)

Version: 2.0

Authors: PPD, Pfizer Inc.
PPD, Medical Research Department,
Cegedim Strategic Data (CSD).

Date: 27-September-2016

[redacted]

PASS information

Title	Cohort Study of Venous Thromboembolism and Other Clinical Endpoints Among Osteoporotic Women Prescribed Bazedoxifene, Bisphosphonates or Raloxifene in Europe.
Protocol number	B1781044
Version identifier of the SAP	2.0
Date of last version of the SAP	27 September 2016
EU Post Authorisation Study (PAS) register number	ENCEPP/SDPP/5395
Active substance	Bazedoxifene
Medicinal product	CONBRIZA
Product reference	EMEA/H/C/000913, EU/1/09/511/001-005
Procedure number	EMEA/H/C/000913/ MEA 012
Marketing Authorisation Holder (MAH)	Pfizer Limited
Joint PASS	No
Research question and objectives	This observational cohort study is being conducted to characterize the risk of selected adverse events of interest among a patient population prescribed bazedoxifene, raloxifene, or a bisphosphonate in usual clinical care outside of a randomized clinical trial setting. The main study objective is to estimate and compare the incidence rates of venous thromboembolism and other clinical endpoints among women receiving bazedoxifene and women receiving either a bisphosphonate or

[redacted]

	raloxifene for treatment of osteoporosis. The main objective of the interim analysis is to determine if, based on the incidence rate of the primary endpoint (venous thromboembolism (VTEs)) and the number of patients accrued in each treatment group, there will be sufficient accrual of data/events to continue the study as planned or if there will be a need to amend the protocol and/or extend the study duration in order to accrue the number of events and patients required to preserve the minimal detectable incidence rate ratios targeted in the protocol.
Countries of study	Spain, Italy
Authors	PPD **, ScD** PPD Epidemiology Pfizer Inc 500 Arcola Road, Collegeville, PA 19426, USA Tel: PPD Fax: PPD Email: PPD PPD **, PhD** Statistician Medical Research Department, Cegedim Strategic Data (CSD)

TABLE OF CONTENTS

1 AMENDMENTS FROM PREVIOUS VERSION(S) 6
2 INTRODUCTION 6
2.1 STUDY DESIGN 7
2.2 STUDY OBJECTIVES 11
3 INTERIM ANALYSES AND FINAL REPORT 11
4 HYPOTHESES AND DECISION RULES 11
4.1 STATISTICAL HYPOTHESES 11
4.2 STATISTICAL DECISION RULES 12
5 ANALYSIS SETS/ POPULATIONS 12
5.1 FULL ANALYSIS SET 12
5.2 SAFETY ANALYSIS SET 12
5.3 OTHER ANALYSIS SET 12
5.4 SUBGROUPS 12
6 ENDPOINTS AND COVARIATES 13
6.1 EFFICACY/ EFFECTIVENESS ENDPOINT(S) 13
6.2 SAFETY ENDPOINTS 13
6.3 OTHER ENDPOINTS 14
7 HANDLING OF MISSING VALUES 14
8 STATISTICAL METHODOLOGY AND STATISTICAL ANALYSES 15
8.1 STATISTICAL METHODS 15
8.1.1 Patient Characteristics 15
8.1.2 Incidence Rates 17
8.1.3 Sensitivity Analysis 18
8.1.4 Time to Event Analysis 19
8.2 STATISTICAL ANALYSES 24
8.2.1 Patient Characteristics 24
8.2.2 Incidence Rates 25
8.2.3 Sensitivity Analysis 25
8.2.4 Time to Event Analysis 26
8.2.5 Summary of Statistical Analysis 26
9 LIST OF TABLES AND TABLE SHELLS 28
10 REFERENCES 28

[redacted]

1 AMENDMENTS FROM PREVIOUS VERSION(S)

Version	Date	Summary of Changes/Comments
Version 1.0	18-November-2013	Not Applicable
Version 2.0	27-September-2016	Administrative changes. Updated SAP format to reflect CT24-GSOP-SD-GL04 Non-interventional (NI) Study Statistical Analysis Plan Template Version 1.0, Date 31-Dec-2013.

2 INTRODUCTION

Note: in this document any text taken directly from the protocol is ***italicised***.

Osteoporosis is characterised by a decrease in bone mass and architectural deterioration of bone tissue. Subtle modifications of bone remodelling, related to abnormalities of bone turnover, can induce a substantial loss of bone over a prolonged period of time. A period of asymptomatic bone loss results in reduced bone strength. When bone loss is sufficient to cause mechanical weakness, fractures may occur spontaneously or as a result of minimal trauma. Osteoporotic fractures cause a substantial clinical and economic burden for society. Age and menopause are the two main determinants of osteoporosis. The cessation of ovarian production of estrogen, at the time of the menopause, results in an accelerated rate of bone loss in women.

Bazedoxifene (BZA) is a third generation non-steroidal selective estrogen receptor modulator (SERM) currently approved in the European Union (EU) for the treatment of postmenopausal osteoporosis in women at increased risk of fracture. BZA is an estrogen receptor ligand that exhibits tissue specific activity: BZA functions as an agonist in the bone and an antagonist in the breast and uterine endometrium. BZA was developed in tablet form and current dosing consists of once daily administration of a 20 mg tablet. BZA (Conbriza) was approved in Europe via the centralized authorization procedure in April 2009. The first EU launch occurred in Spain in September 2010 which was followed by launch in Italy in April 2011.

Currently, one other SERM, raloxifene (Evista®) is marketed in Europe for the treatment and prevention of postmenopausal osteoporosis.

Bisphosphonates are non-hormone compounds that bind to the bone surface and are then taken up by osteoclasts. Bisphosphonates have a profound effect on bone remodeling, and are widely used for the prevention and treatment of osteoporosis.

[redacted]

In clinical trials, BZA-treated women had an increased risk of venous thromboembolic events (VTEs) compared to placebo. VTEs are included in the product label as an important identified risk.

2.1 STUDY DESIGN

This Post Authorization Safety Study (PASS) is an observational, non-interventional epidemiologic cohort study using extracted data from proprietary longitudinal patient databases (LPD) in Italy and Spain. This study is being conducted to further characterize selected adverse events of interest among a patient population prescribed BZA, raloxifene, or a bisphosphonate in usual clinical care outside of a randomized clinical trial setting. This PASS is a post-authorization commitment to the Committee for Medicinal Products for Human Use/ European Medicines Agency (CHMP/EMA).

The LPD databases of Italy and Spain contain all drug prescriptions, diagnoses, demographic data, medical history (event data, risk factors) and other types of data electronically collected by participating general practitioners. The eligible population will be constructed from these anonymised LPD databases.

The total study period is planned for 8 years. This study was launched after BZA became commercially available in the participating countries, Italy and Spain. The subject inclusion/accrual period (also referred to as enrolment period) in the database began at the time of commercial launch of BZA in each country (September 2010 in Spain and April 2011 in Italy) and continued for approximately three years from the start of the study, Table 1. The index date for each patient was the date of the first recorded prescription for BZA or raloxifene or bisphosphonates during the enrolment period. Each woman must have at least 6 months of data prior to index prescription date to be included in the analysis, the baseline period. The study therefore consisted of three distinct periods:

- Pre-enrolment period: 6 months before launch day in a country
- Enrolment period: Approximately 3 years from launch day in a country.
- Follow-up period: 5 years of follow-up per patient from the index date.

Table 1: Study timelines by country

	Pre-enrolment Period 6 Months	**Enrolment Period** 3 Years*	**Follow-Up Period** 5 Years
Spain	April 1st 2010- September 30th 2010	October 1st 2010- April 30th 2014	May 1st 2014- April 30th 2019
Italy	November 1st 2010- April 30th 2011	May 1st 2011- April 30th 2014	May 1st 2014- April 30th 2019

* Due to low enrolment/accrual into the study, the original end of the study accrual period was extended from September 30th 2012 in Spain, and April 30th 2013 in Italy, to April 30th 2014 for both Spain and Italy. Extending the accrual period to April 2014 extended the study timelines by 1 year. The new accrual period for Spain is 43 months (October 1st 2010-April 30th 2014) and the new accrual period for Italy is 36 months (May 1st 2011-April 30th 2014).

[redacted]

Patients will be counted in according to the inclusion criteria.
Follow up for each endpoint will be from the index date to whichever of the following occurs first.

1. Endpoint under investigation
2. Last patient visit or completion of 5 years follow up, whichever occurs first.

Each patient will be followed for identification of selected endpoints for a period of up to five years from their index prescription date for BZA, bisphosphonate or raloxifene. The follow up will continue even if the woman discontinues the index medication or switches to a different medication during this five year follow up period. The follow up duration of an individual patient may be shorter than five years if the woman transfers out of the GP office that wrote the index prescription or if she dies.[1] *In these situations the data will be censored on the date of last visit. This study will use the Cegedim Longitudinal Patient Database (LPD), an electronic medical records database.* However, the Spanish and Italian databases of Cegedim Strategic Data (CSD) do not have recorded events of death and date of death. The discontinued patient's record indicate loss to follow-up for various reasons, such as patients' moving from the areas, patient's change of GPs or other various reasons. The number of patients who are lost to follow-up will be provided. *A minimum of a six month history in the LPD database prior to the index date (at least one GP visit, whatever the motive of the visit to GP) is required for inclusion of population, in order to ensure a certain quality of information.*

Study Population
The study population will be selected based on the following inclusion criteria and all eligible patients enrolled in this study will be included. No exclusion criteria were applied.

Any patient with a prescription for BZA (exposed cohort), raloxifene (comparator cohort) and bisphosphonate (comparator cohort) treatment accompanied with diagnosis of osteoporosis will be eligible when they meet other inclusion criteria. The detailed inclusion criteria will be as follows:

1. *Female*
2. *At least one prescription for BZA, raloxifene, or any bisphosphonate during the study inclusion period*
3. *A recorded diagnosis code of osteoporosis on or within 60 days prior to the index prescription date.*
4. *Age of patients ≥ 45 at the index prescription date*
5. *At least 6 months of follow-up data in the electronic medical record system prior to the date of index prescription.*

[1] The event of death and date of death are not recorded in CSD Italian and Spanish database. The discontinued patient's records indicate loss to follow-up for various reasons.

[redacted]

The applicability of the third criterion will be evaluated through performing a sensitivity analysis comparing those with osteoporosis diagnosis within 60 days and those without diagnosis within 60 days. This has become necessary due to a significant number of patients (> 45%) missing a recorded diagnosis of osteoporosis. If the venous thromboembolism rates of the two groups are comparable (i.e., difference within 20%), this criterion will not be included.

Data Source

This study will use the Cegedim Longitudinal Patient Database (LPD), an electronic medical records database that collects clinical data from primary care physician practices in the following European countries i.e., Spain, Italy, Germany, France, Belgium and UK. (In the UK the database is known as The Health Improvement Network (THIN).) For this study the Cegedim LPD in Italy and Spain will be used.

The Cegedim LPD obtains medical information from proprietary practice management software used during physician office visits to capture clinical data in an electronic medical record system. Physicians use the practice management software developed by Cegedim to maintain electronic medical records of their patients. In each country a panel of physicians using this electronic system volunteered to make available anonymized patient-level information from their practices for clinical research purposes. Since these data are being collected in usual clinical care in a non-interventional way, they reflect routine clinical practice in these countries. The panel of contributing physicians is maintained as a representative sample of the primary care physician population in each country according to age, sex, and geographical distribution. Additionally in most countries (including Spain and Italy), *the patient population is representative of the respective country population according to age and sex distribution, as provided by national statistic authorities.*

Table 2: Doctor and patient populations in the Cegedim LPD by country

	Italy	Spain
Number of physicians in the panel	700	300
Average number of patients who consulted GP at least once in a year	800,000	320,000

The patient data in the LPD form a nationally representative sample. Data have been collected in Italy since 2004 and in Spain since 2006, providing several years of medical history, including comorbidity and concomitant medication use information. Of the patients included in the Spanish and Italian LPD, there are currently approximately 68,000 women with drug-treated osteoporosis. Between 1 and 6% of these patients are taking raloxifene, depending on the region; most of the rest are prescribed a bisphosphonate (Cegedim, data on file).

[redacted]

Data are entered regularly during usual patient care, submitted daily to the Cegedim coordinating center, cleaned and de-identified, and then made available for research.

Anonymized patient data collected from each GP practice includes:

- *Demographic information (age, gender)*
- *Medical history (event dates, diagnoses, risk factors, referrals to specialists)*
- *Therapeutic history (date/length of prescription, molecule/product, dosage)*
- *Additional information (test results, immunizations, height, weight, blood pressure)*

Patient data collected by Cegedim in each country participating in the LPD varies to some extent to accommodate local needs. However, all countries collect data on medical comorbidities and outcomes, prescriptions, demographics, and physician characteristics.

In the Cegedim electronic medical record system diagnosis of clinical events are recorded as diagnostic codes. These codes from the different participating countries will be harmonized based on pre-specified algorithms developed prior to the analysis and listed in the statistical analytic plan.

Cegedim LPD does not collect hospitalization data directly in Italy and Spain. Information on hospitalizations is captured in the patient's general practice file during follow-up visits with the patient's general practice physician following discharge from the hospital. However, this information is not systematically collected and there is no established linkage between the medical records at the general practices and at the hospitals.

Table 2: Data coding conventions in the different countries

	Italy	**Spain**
Drug code dictionary	Farmadati	Vademecum
Therapy classification	Anatomical Therpeutic Chemical (ATC)	ATC
Disease classification	ICD-9	Code ICPC mapped to ICD-9

[redacted]

2.2 STUDY OBJECTIVES

Primary Objective: *To estimate and compare the incidence rates of venous thromboembolism (VTE) among women receiving BZA and women receiving a bisphosphonate for treatment of osteoporosis*.

Secondary Objective 1: *To estimate and compare the incidence rates of VTE among women receiving bazedoxifene and women receiving raloxifene for treatment of osteoporosis.*

Secondary Objective 2: *To estimate and compare the incidence rates of selected clinical endpoints (listed as secondary endpoints in "section 6 Endpoints and Covariates") among women receiving bazedoxifene and women receiving a bisphosphonate for treatment of osteoporosis.*

Secondary Objective 3: *To estimate and compare the incidence rates of selected clinical endpoints (listed as secondary endpoints in "section 6 Endpoints and Covariates") among women receiving bazedoxifene and women receiving raloxifene for treatment of osteoporosis.*

3 INTERIM ANALYSES AND FINAL REPORT

An interim analysis will be conducted once, after four years of study data become available. This interim analysis will include a calculation of crude event rates for the study outcomes in each exposure group. The observed rates of the primary endpoint will be used to determine if the rates are sufficient to continue the study as planned. Based on the results of this interim analysis, the study may be extended in order to accrue the number of events required to preserve the precision of the incidence rate ratios (as defined in the study protocol section 7.1, Sample size calculation)*. There will be no other changes in the study design nor will there be any changes needed to the originally planned analysis as a result of this interim analysis.*

Final analysis will be conducted within 12 months of the end of the study.

4 HYPOTHESES AND DECISION RULES

4.1 STATISTICAL HYPOTHESES

Risk of the primary and secondary endpoints in each treatment cohort is equivalent:

- H_0: Hazardbazedoxifene =Hazardbisphosphonate (Hazard Ratio [HR]=1)
- H_0: Hazardbazedoxifene =Hazardraloxifene (HR=1)

That is, the null hypotheses will be that the rates of the primary endpoints events are the same in patients treated with BZA as in patients treated with bisphosphonate (or raloxifene). The null hypothesis for the analysis of each endpoint will be HR=1; the alternative hypothesis will be HR≠1.

[redacted]

4.2 STATISTICAL DECISION RULES

Incidence rates of the primary and secondary endpoints will be calculated with a 95% Confidence interval (CI). The HR from Cox proportional hazard regression models will be used for comparison of event rates of the primary and secondary endpoints in each exposure (treatment) cohort. Statistical two-sided tests will be used with a significance level of 0.05.

Multiple Comparisons/ Multiplicity

No adjustments for multiple comparisons or multiplicity are planned among endpoints for this Post Authorization Safety Study.

5 ANALYSIS SETS/ POPULATIONS

5.1 FULL ANALYSIS SET

There is no active de novo patient enrolment in this study and all women in the database who meet the inclusion criteria during the enrolment period will be included in the full analysis set.

5.2 SAFETY ANALYSIS SET

The safety analysis set is the same as the full analysis set.

5.3 OTHER ANALYSIS SET

Not Applicable

5.4 SUBGROUPS

Age stratified randomly selected subgroup

In the current dataset, there are 1,823 patients in the bazedoxifene group and 82,041 patients in the bisphosphonate group. This means that there is more than 40 times the number of subjects in the bisphosphonate group compared to the number of subjects in the bazedoxifene group. To limit the number of patients in the bisphosphonate group, an age stratified random subgroup will be selected. Based on the sample size and relative risk estimates provided in the final study protocol (study protocol section 7.1, Sample size calculation), the 1,823 bazedoxifene subjects will be adequately matched by a ratio of 1:6 bisphosphonate subjects. Therefore 1,823*6=10,938 bisphosphonate treated subjects are needed to be randomly selected from the total of 82,041 bisphosphonate treated subjects.

Selection scheme: The age of subjects in the study will be categorized into 4 groups (45-49, 50-59, 60-69, and ≥70). In the bisphosphonate group, the percentage of subjects in each of these age groups is 1.4%, 13.7%, 28.7%, and 56.2% respectively. To maintain the same proportion of subjects in each age group for the bisphosphonate subgroup, there will be 1.4%*10,938=153 bisphosphonate treated subjects randomly selected from the 45-49 age group; 13.7%*10,938=1,499 subjects randomly selected from the 50-59 age group; 28.7%*10,938=3,139 subjects randomly selected from the 60-69 age group; 56.2%*10,938=6,147 subjects randomly selected from ≥ 70 age group.

[redacted]

Incident cases of study endpoints
Women with a history of study end points recorded prior to the index prescription will not be included in the analysis for that specific outcome, but will be analyzed separately as subgroups for each study end point. The analyses of the primary and secondary endpoints will be performed using these incident cases of study endpoints subgroups.

For example, a woman with a history of VTE period prior to the index prescription will be categorized as a prevalent case and will be excluded from the primary VTE analysis. A separate analysis for subgroup of women with a medical history of VTE will be performed and the woman would remain eligible for the analysis of other outcomes as normal. Depending on the number of patients in these subgroups of prevalent cases of each endpoint, this supplemental analysis is likely to have low power to perform any statistical comparison and will be descriptive in nature.

Incident/new user subgroup
In this subgroup analysis, cohorts of incident and not prevalent users of bisphosphonate and raloxifene will be compared to a cohort of incident BZA users. Patients will be defined as incident/new user if they have not had any treatment for osteoporosis in at least 6 months prior to the index date. The index date for each patient is the date of the first recorded prescription for BZA or bisphosphonate or raloxifene during the study enrolment/accrual period.

6 ENDPOINTS AND COVARIATES
The following Primary end points and Secondary end points will be identified with appropriate diagnosis codes. The detailed list of diagnosis codes in each country is available in a separate document.

6.1 EFFICACY/ EFFECTIVENESS ENDPOINT(S)
Not applicable

6.2 SAFETY ENDPOINTS
All the primary and secondary endpoints are safety endpoints (see below).

Primary Endpoints
- Venous thromboembolism (VTE) defined as deep vein thrombosis (DVT), pulmonary embolism (PE), retinal vein thrombosis, and sinus thrombosis

Secondary Endpoints
- Ischemic Stroke
- Thrombotic/Ischemic Cardiac Disorders (including myocardial infarction, myocardial ischemia and coronary occlusion)
- Atrial Fibrillation
- Hypertriglyceridemia

[redacted]

- Biliary Events: Cholecystitis, cholelithiasis
- Clinical Fractures
- Chronic/Acute Renal Failure (including chronic renal insufficiency and end stage renal disease)
- Depression
- Selected ocular events including retinal vascular occlusions, disorders of the globe, iris, ciliary body, retina, eye adnexa and cornea.
- Thyroid disorders
- Aggregate of all malignancies including Breast, Renal, Thyroid, Ovarian, Gastrointestinal tract, Lung cancer
- Malignancies of Breast
- Malignancies of Kidney
- Malignancies of Thyroid
- Malignancies of Ovary
- Malignancies of Gastrointestinal tract
- Malignancies of Lung

6.3 OTHER ENDPOINTS

Not Applicable

7 HANDLING OF MISSING VALUES

For each variable, the number and proportion of patients missing data for each treatment group will be specified. The primary analyses will be done using only the information available in the unmodified dataset. Smoking status, alcohol consumption, height and/or weight (BMI) values are known to be missing for approximately 40-60% patients in the Cegedim LPD based on previous studies. (Age of patients will always be known as it is a selection criterion). Dates are automatically recorded in the Cegedim LPD and doctors can select diseases codes from the pre-defined lists. In all analyses, if a condition (disease or treatment) other than the diagnosis of osteoporosis is not recorded, it will be assumed to be non-existent.

To estimate the effect of missing data, two additional analyses for the primary study endpoint will be conducted. First, the analyses will be repeated utilizing a dataset in which cases with missing values of main study covariates of BMI, smoking status and alcohol status are excluded. The disadvantage to this approach is the reduction in the sample size. Another disadvantage to this approach is that the subjects with missing values may be different than the subjects without missing values (e.g., missing values that are not missing at random), resulting in a non-representative sample after removing the cases with missing values. Second, multiple imputations for the missing values of BMI, smoking status, alcohol status will be used in the second set of analyses and compared to complete case analyses to understand the potential impact of missing data on results [2].

Outlier

[redacted]

Extreme values for weight and height will be searched for. If a value seems implausible, as regard to gender or ratio weight/height, it will be put as missing.

8 STATISTICAL METHODOLOGY AND STATISTICAL ANALYSES

8.1 STATISTICAL METHODS

Analyses will be conducted separately for each country and the final report will consist of three parts: one part for Italy, one part for Spain, and the final part with pooled data of both countries. Pooling of data for the combined results of Italy and Spain will be done at the patient level. ICD9 codes of the Italian data will be mapped unto ICPE codes of the data from Spain. ICD9 codes that do not map directly to ICPE will be noted and the frequency of patients with outcomes in these unmapped categories will be reported. Given the very low enrolment from Italy (approximately 40 patients, vs. approximately 1700 from Spain) very few patients/outcomes from Italy are likely to be reported as unmapped. Assessment and adjustment for confounding will be done separately for each country and also done separately for the pooled analyses and follow the same methodology in the individual and pooled analyses. All statistical analyses will be performed using the SAS software 9.2, running with SAS Enterprise Guide 4.3, on a Windows Server.

The qualitative variables will be described by using usual statistics: population size and percentages (relative to the non-missing data) for each modality. Quantitative variables will be described with the number of observed data, mean, standard deviation (SD), median, and ranges. Two-sided 95% confidence intervals will be presented for relevant variables.

8.1.1 PATIENT CHARACTERISTICS

The following characteristics will be reported for all treatment cohorts at baseline.

- Demographic Characteristics: Age, BMI, Smoking

  Age (years) at index date will be calculated as Age = year (index date)-year (birth).

  Four age groups are defined (in years): 45-49, 50-59, 60-69 and 70 and over.

  BMI (Body Mass Index) will be calculated for each patient as the weight in kg divided by square of height in meters. BMI will be calculated using the following formulas:

$$BMI\ (kg/m^2) = Weight\ (kg)/(Height\ (m))^2$$

Or

$$BMI\ (kg/m^2) = 10000 \times Weight\ (kg)/(Height\ (cm))^2$$

[redacted]

BMI will be categorized into four groups as follows:

- Underweight: BMI <18.5 kg/m²
- Normal range: 18.5 ≤ BMI <25 kg/m²
- Overweight: 25 ≤ BMI <30 kg/m²
- Obese : BMI ≥30 kg/m²

- Smoking will be defined as three categories; Current Smokers, Former Smokers and Never Smokers. A 'Current Smoker' will be defined as a patient who is currently smoking at the index prescription date, a 'Former Smoker' as a patient that has a past history of smoking but has stopped prior to the index prescription date. A 'Never Smoker' will be defined as a patient that has no history of smoking (The reported frequency of smoking may be very low because smoking status is under recorded in Italian and Spain databases)

- Alcohol use. Alcohol use will be defined as three categories; Current Drinkers, Former Drinkers and Never Drinkers. A 'Current Drinker' will be defined as a patient who is currently drinking alcohol as at the index prescription date, a 'Former Smoker' as a patient that has a past history of drinking alcohol but has stopped drinking prior to the index prescription date. A 'Never Drinker' will be defined as a patient that has no history of drinking alcohol. (The reported frequency of alcohol use may be very low because alcohol use is under recorded in Italian and Spain databases).

- Use of relevant concomitant medications, e.g., hormone replacement therapy. (A concomitant medication is any medication used during BZA/bisphosphonate/raloxifene treatment defined as the total treatment duration between the first and the last prescription of BZA/bisphosphonate/raloxifene.)

- Medical History VTE (Venous Thromboembolic Events), Diabetes, Hypertension, and Malignancies (including Breast, Renal, Thyroid, Ovarian, Gastrointestinal tract, Lung) and aggregate of all malignancies.

  Venous Thromboembolic Events (VTE) are defined as deep vein thrombosis, pulmonary embolism, retinal vein thrombosis, and venous sinus thrombosis. This objective is to compare this incidence in real life settings with validated diagnosis codes, ICD9 in Italy and International Classification of Primary Care (ICPC) codes in Spain.

The availability of medical history of all relevant medical diagnosis for patients will be checked in both the Spanish and the Italian database and presented if available. The variables of medical history (diabetes, hypertension, VTE and Malignancies) will be presented in classes (Yes/No) before index date. (All the detailed diagnosis codes of each endpoint have been validated by Pfizer and are available in a separate reference document). For all medical history and smoking status, if the relevant information is not found in the database, it will be assumed to be negative.

[redacted]

8.1.2 INCIDENCE RATES

Incidence rates will be calculated for all endpoints in all cohort groups; incidence rate ratios and 95% confidence intervals will be calculated as specified previously.

The incidence rate for all end points will be calculated using both the cumulative incidence approach and an incidence density (person-time) approach. The cumulative incidence approach will be used as the primary analysis and it will be calculated as in the following formula:

Incidence rate = (Total new cases during follow-up period / Total persons at risk during follow-up period)

For this study, the numerator will be defined as the total number of patients with a diagnosis of primary end points or secondary end points during follow-up period. The denominator will be the total number of all patients at risk in cohorts groups during follow-up period.

Secondly, an incidence density approach calculates time at risk in person-years, measured from first prescription to the end of the observation period. The incidence rate will be reported as per 1,000 person–years of observation as shown in the following formula:

*Incidence Rate = (Total new cases during follow up period / Total person years during follow-up period)*1000*

The numerator will be defined as the total number patients with end points of interest during follow-up period. The denominator will be defined as the sum of person years of patients during the follow up period. For patients who have incidence of primary or secondary end points, person time will equal the sum of days from their first prescription of treatment (index prescription) until the date of the incident condition (diagnosis of primary or secondary end points). For patients who do not experience an incidence of primary or secondary end points, person time will equal the sum of days from their first prescription of treatment (index prescription) to the last visit date or completion of 5 years follow up at which all patients will be censored, whichever occurs first.

In the determination of person years, the date of first prescription of treatment will be used as a surrogate for time of first exposure to treatment, since the date when the patient had their first treatment for osteoporosis may not be consistently ascertainable.The same definition of person-time ('ever' exposed approach accounting for all person-times of exposure) will be used for all the study endpoints given that they are all chronic diseases or malignancies with potentially long latency periods. If no primary end point occurs during the follow-up period, the duration will be calculated from the start of treatment date to the last patient visit or the end of study period, whichever comes first.

Both methods of calculating incidence rates will be used and the number of patients for whom the date of first treatment for osteoporosis could not be ascertained will be provided as part of the sensitivity analysis of the incidence density approach.

Stratifications by risks of interest (e.g. age, BMI, or selected medical history) will be performed for each outcome and descriptive summaries will be reported. For example, VTE incidence will be stratified according to age as well as by BMI. Other stratifications by medical co-morbidities and other variables of interest will also be performed as appropriate.

Given the rarity of certain malignancies in this population, the incidence rates observed in this study may be small and imprecise. To better characterize the rates obtained in this study, an additional analysis using external reference data for the malignancies will be conducted, i.e. a standardized incidence ratio (SIR). The observed frequency of malignancies in all three treatment groups will be compared with the expected frequency calculated from an age and gender matched population-based reference group (e.g., GloboCan). This SIR of observed versus expected events and 95% confidence interval, will be reported.

8.1.3 SENSITIVITY ANALYSIS

For each prescription, duration of drug exposure will begin to accrue from the date of the prescription. The duration of drug exposure will be categorized into:

- ***Current exposure:*** *defined as the duration of the prescription*
- ***Recent exposure:*** *defined as the 60-day period following the end of the duration of prescription*
- ***Past exposure:*** *defined as any follow up time accrued after 60 days following the end of the duration of the prescription.*

Given the discontinuous nature of osteoporosis medication use, it is likely that a large number of women in this study will accrue 'current exposure', 'recent exposure' and 'past exposure' exposure time to multiple osteoporosis drugs during the 5-year period of follow up.

A sensitivity analysis will be conducted that compares using a 30-day and a 90-day definition of recent exposure with using the 60-day definition.

Exposure to the drugs under study will be ascertained by searching the prescription records for the relevant product codes. For each prescription the duration of exposure will be estimated based on the dose and quantity of the prescribed medication. For patients that switch treatment, the as-treated (AT) approach will be used in calculating the exposure time (see section below on patients switching treatment for more details).

There will be rules for handling missing values of dose or quantity prescribed in Italian and Spanish LPD databases. In LPD databases, the number of days of supply or daily

[redacted]

dose is not recorded by some GPs in Spain and Italy. If duration is not available, the number of days of supply will be calculated by dividing total number of pills (number of prescribed box times number of pills in a box) by the daily dose. If duration and daily dose are not available, the number of days' supply will be calculated by dividing total number of pills (number of prescribed box multiplied by the number of pills in a box) by the standard dose/recommended dose.

Drug exposure will also be categorized with respect to the cumulative exposure to each drug (e.g., 0-days, 1-30-days, 31-90-days, 91-180-days, 181-365-days, >365-days) to evaluate whether or not the incidence (hazard) of a specific adverse event varies with increased duration of use.

8.1.4 TIME TO EVENT ANALYSIS

A time to event analysis to compare the risk between treatment cohorts, i.e., the BZA, raloxifene and bisphosphonate treated groups, will be conducted for the primary end point (VTE) and other secondary endpoints if sufficient data are available,

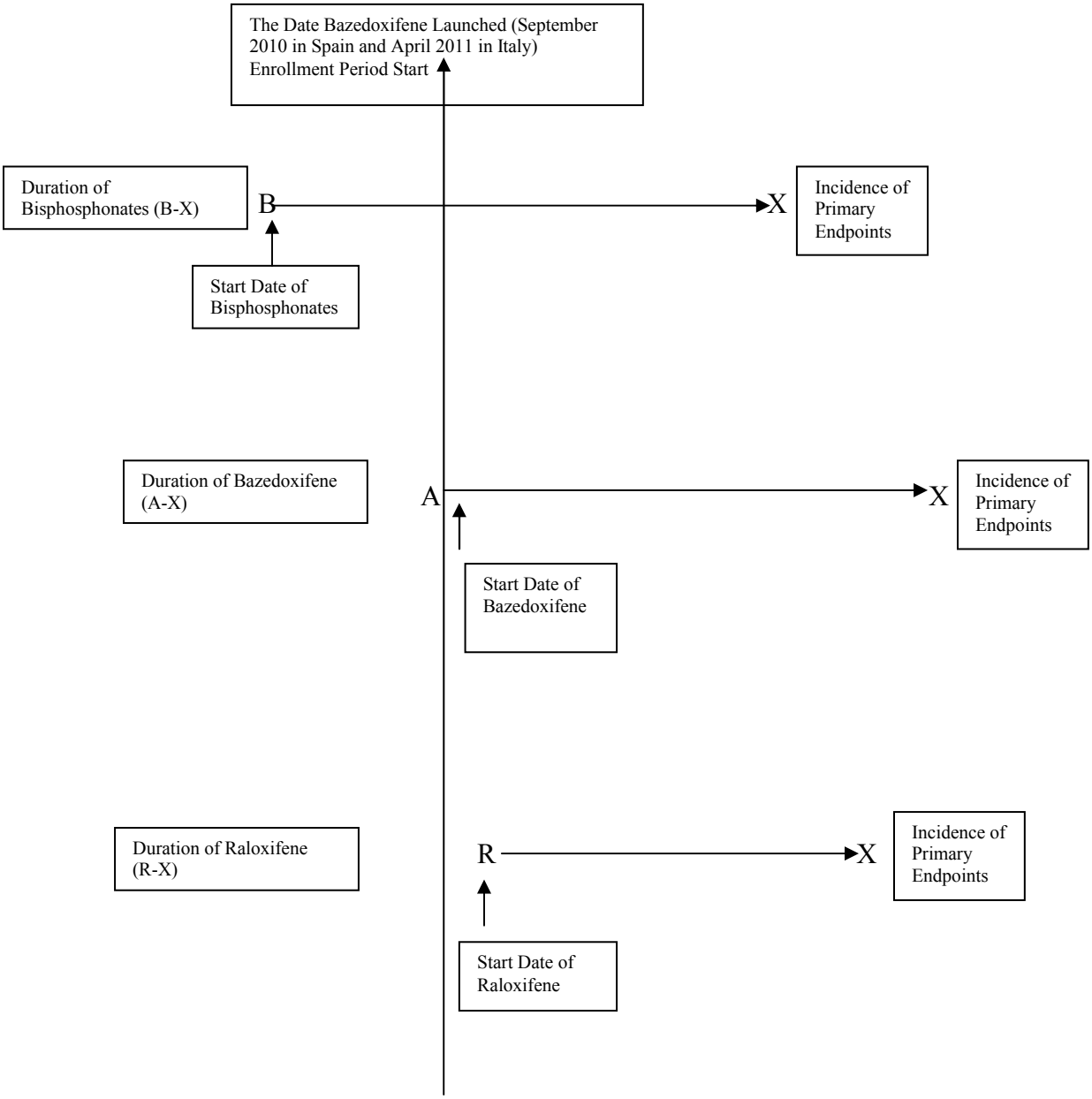
The time to an event will be calculated as the time from index prescription date to the incidence date of primary end points. If a patient does not have a primary end point, the treatment duration will be based upon the index prescription date to the last visit or end of follow-up period, whichever comes first.

In order to ensure the exact duration from the start of treatment to the incidence of primary endpoints, a subgroup of newly prescribed patients who were diagnosed with osteoporosis during the enrolment period will be used first.

In the very unlikely event that the number of bisphononate or raloxifene patients are less than the number of BZA patients, the analysis will be extended to include bisphosphate and raloxifene patients whose treatment began prior to the enrolment period. Considering the fact that bisphosphonates and raloxifene were launched before BZA (See the figure below), some patients could have already gotten treatment of bisphosphonates and raloxifene prior to the enrolment period (i.e., before BZA was launched). For such patients the start date will be the first date of treatment for bisphophonates or raloxifene, available within the database.

[redacted]

Figure 1.Calculation of duration (from start date to the incidence of primary endpoint) for the treatment cohorts



The time to event analysis model will also be used to test for differences in time to event while adjusting for covariates of interest.

This model will take into account the duration to the incidence of the endpoint and control for imbalances in outcome risk factors, as determined based on the baseline descriptive analysis.

[redacted]

The following factors possibly related to the incidence of the primary end point will be considered.

- Treatment group (BZA, raloxifene, and bisphosphonate)
- Age
- BMI
- Reported history of smoking (If the reported frequency of smoking status is low this will be excluded)
- Reported Medical History of Diabetes, Hypertension, Malignancies
- Time since diagnosis of osteoporosis
- Previous treatment for osteoporosis (before index date) or Number of different treatments for osteoporosis in the entire patient history available
- History of DVT or PE
- Previous history of major lower extremity (hip, knee) arthroplasty, arthroscopic surgery (information on arthroscopic surgery not available in Italian database)
- General surgery, major orthopedic surgery (<1month prior to incidence of VTE)
- (Information on general surgery not available in Italian database)
- Multiple trauma (<1month prior to incidence of VTE)
- Hip, pelvis, or leg fracture (<1month prior to incidence of VTE)
- Stroke (<1month prior to incidence of VTE)
- Acute spinal cord injury (paralysis) (<1month prior to incidence of VTE)
- Hormone replacement therapy (limited information available in both databases)

A Cox model will be used in order to take into account the time varying component.
The time from the beginning of the treatment (one among three types, BZA, raloxifene, and bisphosphonate) to the occurrence of primary end point is fitted in this model. The Proportional Hazards (PH) assumption will be tested for each covariable. If the PH assumption does not hold for some covariates, the use of time-varying HR for those covariates will be explored. If a categorical variable with more than 2 levels does not satisfy this PH assumption, collapsing some categories will be attempted to satisfy the PH assumption. Correlations between variables will also be evaluated in order to find independent factors and a final model that is clinically relevant.

First, the univariate analysis will be carried out on each covariable in order to determine the explanatory variables to be included in the final multivariate model. All significant variables with a p-value ≤ 0.20 in the univariate model will be taken into account in the multivariate Cox model.

If > 80% of the important risk factors for VTE listed above does not satisfy the PH assumption, a logistic regression analysis would be performed instead of Cox model to otherwise determine if the covariable is related to the incidence primary endpoint (VTE). The dependent variable would be the incidence of the primary end point during the follow-up period categorized as a binary variable. The 2 indicators of treatment group (BZA vs. bisphosphonates, raloxifene, will be forced in the multivariate model (Cox model or logistic model) to analyze incidence in each class of treatment, adjusted on

[redacted]

significant factors. Hazard ratios (HR) (odds ratios (OR) in case of logistic regression) with 95% Confidence intervals (CIs) and p-values will be presented.

Patients Switching Medications

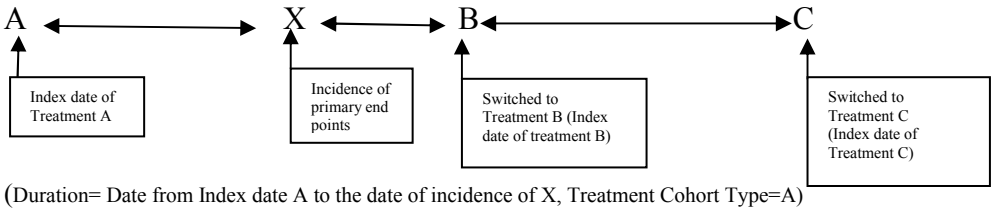
In the treatment of chronic conditions such as osteoporosis, patients may discontinue their study medication (BZA/bisphosphonate/raloxifene) or change study medications. Such exposure changes are often due to treatment failure or perceived/real side effects. The stronger such non-adherence is associated with the outcome; the more an as-treated (AT) analysis, which censors at the point of discontinuation, will be biased. A cumulative risk (CR) (similar to intention-to-treat analyses in randomized trials) analysis follows all patients for a fixed time period, carrying forward the initial exposure status and disregarding any changes in treatment status over time. Because this analysis disregards informative non-adherence, it will not suffer bias as a consequence of censoring, but it will suffer bias as a consequence of exposure misclassification. Such misclassification increases with a longer follow-up period and a shorter average time to discontinuation. In most cases, though not all, such misclassification will bias effects towards the null, similar to intention-to-treat analyses in randomized trials. Viewed separately, the AT and CR analyses trade biases, but together they give a range of plausible effect estimates. Adjusting for non-adherence in an analysis of a drug effect requires information about the predictors of treatment discontinuation, which is not available with sufficient accuracy in secondary data such as that used in this study [1].

In this study the proportion and demographic and clinical characteristics of patients that switch medications between the study drugs will be reported alongside patients that do not switch medications. Both the AT and CR analytic approaches will be used for all study endpoints and the results reported. In the CR method, if a patient switches from one exposure (study drug) cohort group to another, they will have several treatment durations and will belong to multiple cohort groups but will be analyzed according to their original exposure cohort group (i.e., CR/intent to treat approach). In the AT approach, all exposure cohort groups will be included as a time-dependent variable. With the exposure cohort group as a time-dependent variable, each event will be assigned to the treatment being taken at the time of occurrence. The duration of treatment for each cohort group will be from the date of the first prescription to the date of the event or to the end of treatment, whichever occurs first. The covariates of patients that switched treatment at the time of the event or at the end of follow-up would be considered in this model.

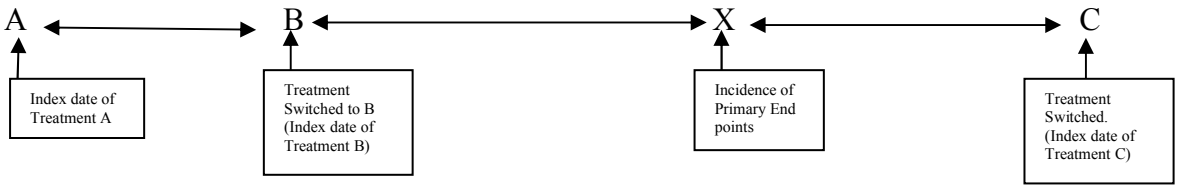
The following diagrams shows some example of calculating duration for patients with switched treatments with both the AT and CR approaches.

[redacted]

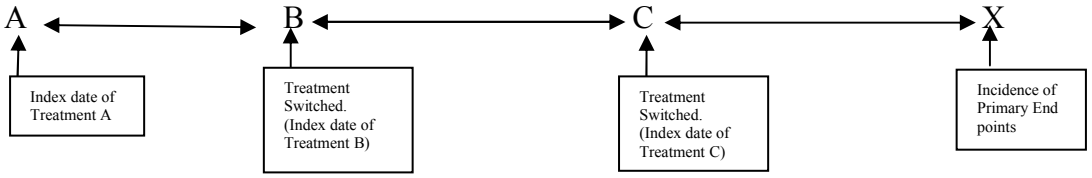
Figure 2a. Calculating the duration of treatment for patients that switched treatment using the as-treated (AT) approach.



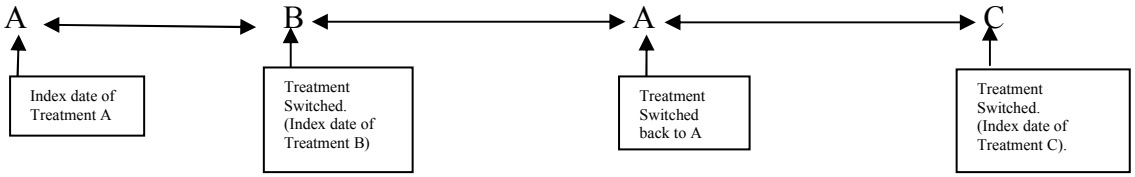
(Duration= Date from Index date A to the date of incidence of X, Treatment Cohort Type=A)



(Duration= Date from Index date of Treatment B to the date of incidence of X, Treatment Cohort Type=B)
(Time from A to B will be counted towards treatment A)



(Duration= Date from Index date of Treatment C to the date of incidence of X, Treatment Cohort Type=C) The person time will be apportioned as a time varying exposure such that this person accrues person-time for exposure A until point B, then accrues toward exposure B until point C, then accrues to exposure C until point X. The event, X, is only associated with exposure C assuming no induction period, but person time from drugs A and B would contribute to respective exposure groups.

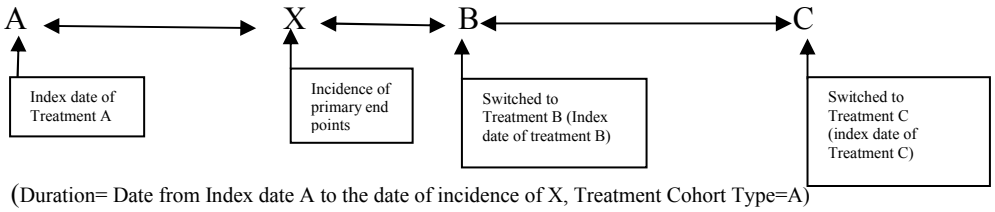


(Duration of patient without incidence = Date from Index date of Treatment C to the last date of follow-up or study end date, Treatment Cohort Type=C)
(This person accrues person-time for exposure A until point B, then accrues person-time for exposure B until the next point A, then accrues person-time for exposure A until point C)

A, B, C: Treatment cohort types. X= Incidence date of primary end point

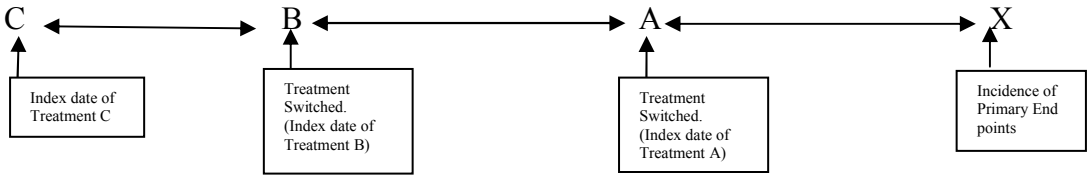
[redacted]

Figure 2b. Calculating the duration of treatment for patients that switched treatments using the cumulative risk (CR) approach.

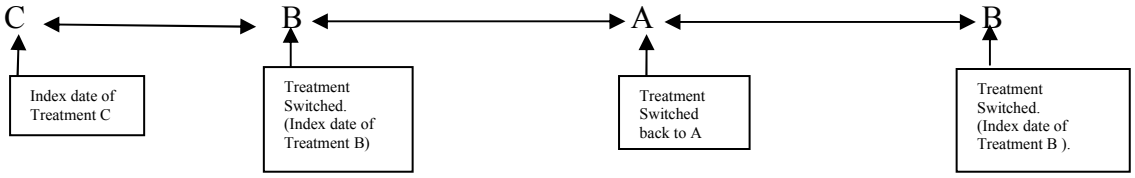


(Duration= Date from Index date A to the date of incidence of X, Treatment Cohort Type=A)

(Duration= Date from Index date of Treatment B to the date of incidence of X, Treatment Cohort Type=B)



(Duration= Date from Index date of Treatment C to the date of incidence of X, Treatment Cohort Type=C)



(Duration of patient without incidence = Date from Index date of Treatment C to the last date of follow-up or study end date, Treatment Cohort Type=C)

A, B, C: Treatment cohort types. X= Incidence date of primary end point

8.2 STATISTICAL ANALYSES

8.2.1 PATIENT CHARACTERISTICS

Descriptive statistics on available patient characteristics (e.g. age, BMI, smoking, history of relevant medical diagnoses) will be reported for all exposure groups at index date.
A summary table will present and compare all the patients' demographic data at index date (BZA/raloxifene/bisphosphonates).

- Age (years): derived by difference between the year of time of interest and year of birth.
- Weight (kg): recorded at the given date, or if not available, at the most recent date before or it
- Height (cm): the value at the given date or, if not available, at the most recent date before or after it.

[redacted]

- Body Mass Index (BMI) (kg/m²), calculated from weight and height as defined previously. BMI will be described in quantitative measure and in categorical data, as usual
- Alcohol use
- Smoking status
- Time since diagnosis of osteoporosis: derived by the difference between the index date and the diagnosis date.
- Previous treatment for osteoporosis (before index date) (yes versus no) or Number of different treatments for osteoporosis in the entire patient history available
- Previous history of VTE (Venous Thromboembolic Events) (before index date)
- Previous history of major lower extremity (hip, knee) arthroplasty, arthroscopic surgery (information on arthroscopic surgery not available in Italian database)
- General surgery, major orthopedic surgery (<1month prior to incidence of VTE) (information on general surgery not available in Italian database)
- Multiple trauma, (<1month prior to incidence of VTE)
- Hip, pelvis, or leg fracture (<1month prior to incidence of VTE)
- Stroke (<1month prior to incidence of VTE)
- Acute spinal cord injury (paralysis) (<1month prior to incidence of VTE)
- Hormone replacement therapy (limited information available in both databases)

If the diagnoses for any of the conditions listed in this section (8.2.1) are not found in the database, they will be assumed to be absent.

8.2.2 INCIDENCE RATES

Incident rates of all primary and secondary endpoints for all three cohort groups (BZA/raloxifene/bisphosphonates) and 95% CI will be presented. The Hazard Ratio (HR) from Cox proportional hazard regression models will be used for comparison of event rates of the primary and secondary endpoints in BZA vs. bisphosphonate and BZA vs. raloxifene. A likelihood ratio test will test the null hypothesis of no difference in hazards in each pair-wise comparison. Statistical tests will use an alpha-level of 0.05.

The cumulative drug exposure of each treatment cohort will be categorized (1-30days/31-90 days/91-180 days) and used to evaluate whether or not the incidence rate of the primary and secondary end points varies with increased duration of treatment.

8.2.3 SENSITIVITY ANALYSIS

Sensitivity analysis will be conducted to compare the different definitions of exposure categories (60 days window vs. 30 day window, 90 day window vs. 30 day window to define recent exposure).

[redacted]

8.2.4 TIME TO EVENT ANALYSIS

Time to event analysis will be conducted as a secondary analysis and the following factors possibly related to incidence of end points will be considered at the date of Incidence or the last visit before incidence: age, BMI, smoking (depending on availability of data), history of diabetes, history of hypertension, history of malignancies, history of VTE, and some other co-morbidities. Univariate Cox models (if applicable, depending on the PH assumptions for each factor, otherwise logistic regression model) will be performed for each factor and then, the analysis will be complemented by a multivariate Cox model (if applicable, depending on the PH assumptions for each factor) or a logistic regression. Hazard ratios (HR) or odds-ratios with 95% CI and p-values will be presented for significant factors at the 0.05 level.

8.2.5 SUMMARY OF STATISTICAL ANALYSIS

Analyses for Primary Objective and Secondary Objective 1

Comparator: BZA vs. bisphosphonate (primary objective), BZA vs. raloxifene (secondary objective 1).

Study endpoint: VTE.

Populations: a. Full BZA full raloxifene, random age stratified bisphosphonate subgroup
b. Incident user subgroup.

Statistical Method:

i. Incidence rate, cumulative method (*(Total new cases during follow-up period / Total persons at risk during follow-up period)* with stratification by risks of interest e.g. Age, BMI.
ii. Univariate and multivariate incidence rate by incidence density *(Total new cases during follow up period / Total person years during follow-up period)*1000) method* using Cox regression model and a Likelihood Ratio Test for testing Hazard Ratio (HR) =1 (ratio of hazard in BZA population vs. bisphosphonate/raloxifene population) with covariates associated with VTE outcome and exposure (Drug groups).

Sensitivity analyses:

i. Repeat analysis will be conducted that compares using a 30-day and a 90-day definition of recent exposure with using the 60-day definition.
ii. Missing Data: First, the analyses will be repeated utilizing a complete dataset in which cases with missing values of main study covariates of BMI, smoking status and alcohol status have been deleted. Second, multiple imputation for the missing values of BMI, smoking status, alcohol status will be used in the second set of analyses.

Multiple imputation procedure:

[redacted]

1) Estimate distribution of BMI, smoking status and alcohol status according to patients whose data are available.
2) For the patients who had missing values those variables draw values of the variables from the distribution from 1).
3) With the imputed dataset (including none-missing data), estimate HR as described.
4) Repeat 2)-3) for *m* times, and record HR of each time.
5) Calculate mean of *m* estimates of HR.
6) Adjust variance of HR with the following formula, and use the variance to calculate confidence interval.

$$V = \left(1 + \frac{1}{M}\right) V_B + V_W.$$

V_B is the between imputation variance, and V_W is within imputation variance.

Analyses for Secondary Objective 2 and Secondary Objective 3
Comparator: BZA vs. bisphosphonate (secondary objective 2), BZA vs. raloxifene (secondary objective 3).
Study endpoint: Secondary endpoints.

Population: a. Full BZA, full raloxifene, random age stratified bisphosphonate subgroup.
b. Incident user subgroup

Statistical Method:

i. Incidence rate, cumulative method (*(Total new cases during follow-up period / Total persons at risk during follow-up period)* with stratification by risks of interest e.g. Age, BMI.
ii. Univariate and multivariate (if covariates available) incidence rate by incidence density *(Total new cases during follow up period / Total person years during follow-up period)*1000) method* using Cox regression model and using a Likelihood Ratio Test for testing Hazard Ratio (HR) =1 (ratio of hazard in BZA population vs. bisphosphonate/raloxifene population) with covariates associated with secondary endpoints and exposure (Drug groups).

Sensitivity analyses:

i. Repeat analysis will be conducted that compares using a 30-day and a 90-day definition of recent exposure with using the 60-day definition.
ii. Missing Data: First, the analyses will be repeated utilizing a complete dataset in which cases with missing values of main study covariates of BMI, smoking status and alcohol status have been deleted. Second, multiple imputation for the missing values of BMI, smoking status, alcohol status will be used in the second set of analyses.

[redacted]

Multiple imputation procedure:

1) Estimate distribution of BMI, smoking status and alcohol status according to patients whose data are available.
2) For the patients who had missing values those variables draw values of the variables from the distribution from 1).
3) With the imputed dataset (including none-missing data), estimate HR as described.
4) Repeat 2)-3) for *m* times, and record HR of each time.
5) Calculate mean of *m* estimates of HR.
6) Adjust variance of HR with the following formula, and use the variance to calculate confidence interval.

$$V = \left(1 + \frac{1}{M}\right) V_B + V_W.$$

V_B is the between imputation variance, and V_W is within imputation variance.

9 LIST OF TABLES AND TABLE SHELLS

Not Applicable

10 REFERENCES

1. Schneeweiss S. A basic study design for expedited safety signal evaluation based on electronic healthcare data. Pharmacoepidemiol Drug Saf. 2010 Aug;19(8):858-68

2. Sterne JAC, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ4 : British Medical Journal* 2009;338:b2393. doi:10.1136/bmj.b2393

[redacted]

Document Approval Record

Document Name:	BZA Study Statistical Analysis Plan (SAP)_SEPT_2016_Final
Document Title:	BZA Study Statistical Analysis Plan (SAP)_SEPT_2016_Final

Signed By:	Date(GMT)	Signing Capacity
PPD	04-Oct-2016 14:12:22	Manager Approval